DOCUMENT RESUME

ED 191 467 IR 008 722

AUTHOR Kim, Chai: Yerkey, A. Neil
TITLE Impact of Other Disciplines on Information Retrieval and Librarianship.
PUB DATE 80
NOTE 12p.: Paper presented at the Mid-Year Meeting of the American Society for Information Science (Pittsburgh, PA, May 14-17, 1980). Best copy available.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Citations (References): *Information Retrieval: *Interdisciplinary Approach: *Library Science: *Relationship
IDENTIFIERS *Bibliometrics

ABSTRACT

The central focus of this research was to analyze the impact of ideas originating from within librarianship and information retrieval and those from other disciplines. Twelve most cited journals in librarianship and information retrieval (source specialties) were chosen for citation analysis. From these journals, 40 articles were randomly selected from the last complete year and their references recorded, resulting in 1,159 citations. The analysis consisted of measurements of self-citation, cross-fertilization (proportion of citations made by one to the other source specialty), and openness (the degree to which a specialty is open to the ideas of other disciplines). Results showed that about half of all citations were self-citations, which is lower than the 70 percent average citation rate found in other scientific disciplines. The two source citations did not cite each other to any great extent. They drew about half of their citations from other disciplines; however, they seemed to be influenced by two different sets of disciplines. Suggested implications are that information retrieval and librarianship are interdisciplinary in nature. Required knowledge for problem-solution may be drawn from any pertinent disciplines as the need arises, rather than as a result of internal intellectual, systematic development. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED191467

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IMPACT OF OTHER DISCIPLINES ON INFORMATION RETRIEVAL AND LIBRARIANSHIP

BY

Chai Kim

and

A. Neil Yerkey

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Chai Kim

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR008722

IMPACT OF OTHER DISCIPLINES ON INFORMATION RETRIEVAL AND LIBRARIANSHIP

Chai Kim
and
A. Neil Yerkey
State University of New York
Buffalo, New York

A scientific paradigm is a set of widely shared understandings that specify a discipline's priorities in selecting research topics, methodologies, and techniques. Two major theses have been dominant in identifying the factors that determine the paradigm of a discipline. "Internalists" view the paradigm as relatively autonomous and independent of external factors. By contrast, "externalists" maintain that the factors of a paradigm originate from outside the discipline altogether.

The central focus of this research was to analyze the impact of ideas originating from within librarianship and information retrieval and those from without, i.e., other disciplines. Based on the "impact factors" of journals reported in SSCI, Journal Citation Reports (Philadelphia, ISI, 1979), twelve most cited journals in librarianship and information retrieval (source specialities) were chosen. From these journals, 40 articles were randomly selected from 1978 and their references recorded. This resulted in 1159 citations for analysis.

The analysis consisted of measurements of self-citation, cross-fertilization (proportion of citations made by one to the other source specialty), and openness (the degree to which a specialty is open to the ideas of other disciplines).

About half of all citations were self-citations, which is lower than the 70% citation rate attributed to science as a whole. The two source specialties did not cite each other to any great extent, in spite of their historical affinity. They drew about half of their citations from other disciplines; however, they seemed to be influenced by two different sets of disciplines.

Suggested implications are that information retrieval and librarianship are interdisciplinary in nature. Their paradigmatic content may be dictated by various problems which they face and feel compelled to solve rather than as a result of internal intellectual, systematic development of knowledge.

IMPACT OF OTHER DISCIPLINES ON INFORMATION RETRIEVAL AND LIBRARIANSHIP

Chai Kim and A. Neil Yerkey
State University of New York at Buffalo

I. Introduction

Throughout its history, information retrieval has had a close relationship with traditional librarianship. Early works by Luhn, Fairthorne, Perry, Kent, Mooers, and Taube, among others, laid the foundations for today's information retrieval technology. People working in information retrieval have maintained a dialogue with librarians and network/computer specialists. From the former, they learned techniques of bibliographic control, and from the latter how to design large databases. This integration of bibliographic control and computer technology provided much of the fundamental knowledge needed for the design and operation of large bibliographic utilities and databases, such as those of OCLC, Inc., System Development Corporation, Lockheed Information Systems, and Bibliographic Retrieval Service, Inc. These in turn helped bring about "the online revolution in libraries."

Despite the importance of librarianship-information retrieval interface, few systematic efforts have been made to analyze the relationship between the two. This paper analyzes the citation patterns to explore the relationship between librarianship and information retrieval. It will attempt to characterize the dynamics of the relationship to help answer such questions as: Is information retrieval technology drawing from the reservoir of knowledge of librarianship, or vice versa? Are the present concerns of the two so intellectually similar that they draw necessary knowledge primarily from each other and not significantly from other disciplines? What is the extent of impact of other disciplines on librarianship and information retrieval?

II. Citation Analysis as a Tool for Studying Scientific Activities

Almost 40 years before citation analysis really came into its own with the birth of the Institute for Scientific Information in 1963, Gross and Gross used citation counts to evaluate journals (Gross, et al, 1927). Since then many researchers have conducted similar research in basic and applied sciences. Some concluded that citation frequency is a valid indicator of the use, and therefore the value, of journals (Garfield, 1972; Hafner, 1976; Virgo, 1977; Windsor, 1973). Others argued that citations may not reflect true usage and thus may be poor indicators of the quality of cited items (Martyn, 1975; Scales, 1976; Subramanyan, 1975).

Despite the controversy, citation analysis has emerged as an important tool for studying science from a sociological perspective. Data from citation analysis strengthened that research which attempted to provide a sociological account for the growth of science and which explored communication among scientists within and between disciplines. In addition, data from citation analysis have been used to evaluate various factors affecting social stratification of science, such as the reward system in science--tenure and promotion, the process of refereeing papers for publication, etc. Recent books by Cole and Cole, Crane, and Meadow

in addition to those papers reviewed in the following illustrate the point (Cole, et al, 1973; Crane, 1972; Meadow, 1974).

Three groups of studies using citation analysis seem to be relevant to this paper. The first group includes those which examined the specialty structure of a discipline in terms of its topical emphasis and the growth of knowledge. Others in this group attempted to forecast the structure of scientific activity (Mekhtiev, et al, 1977; Small, 1973; Borenius, et al, 1972).

The second group includes those which studied the literature use of a discipline as a function of time. In order to examine the relationship between the frequency of use of a paper and its age, publication dates of cited papers have been plotted, producing citation decay curve for different disciplines (Burton, et al, 1977; Price, 1965; Weiss, 1960).

The third group includes those which deal with formal communication among scientists within a discipline and/or between disciplines. A frequently adopted operational definition of formal communication is the degree to which a given group of scientists on a particular discipline cites publications from other scientists working in the same or different disciplines (Holton, 1962; Meadows, 1967; Parker, et al, 1967; Swanson, 1966). This gives an indication of the extent a given discipline is open or closed to the idea of other disciplines. The primary interest in this group of studies is to investigate the extent to which cross-disciplinary diffusion of ideas is a factor affecting the development of a discipline.

Findings of the studies of the first two groups are based on several assumptions. One is that the value of a scientific paper may be measured by the use it receives as indicated by citation frequencies. Another assumption is that primary sources such as journals--the typical source of citation statistics--are representative of the entire literature of the field being examined.

While these assumptions remain problematic to some researchers, it is important to realize that following up citations on relevant paper is still an important source of obtaining information for scientists (American Institute of Physics, 1970). Citation analysis does provide a valid measurement (certainly not the only measurement) of data for the type of research reviewed and conducted in this paper.

III. Models of Cross-fertilization of Ideas between Disciplines

Results of the research using citation analysis seem to derive interpretive powers from two major models which have been dominent in sociology of knowledge. They are the so-called "internalist" and "externalist" models.

T. Kuhn, J. Cole, S. Cole, W. Hagstrom, and other internalists view the paradigm as relatively autonomous and independent of external factors. To them, the mode of change in science is derived from within rather than from without a discipline. Scientists within a discipline are likely to be too busy communicating among themselves to be involved in the exchange of ideas from outside. Knowledge generated from within a discipline is viewed to be all that is necessary to solve the problems of the discipline.

Largely on its own, a discipline identifies problems and the methods for solutions. It sets priorities on different research problems relatively independent of the impact of other disciplines. Hence, a discipline is viewed as being composed of clearly defined areas whose communication is more or less restricted to the discipline to which they belong.

By contrast, J.D. Bernal, R.K. Merton, D. Crane and other externalists maintain that the factors of a paradigm originate from outside the discipline altogether. They explain the content of a paradigm and its change from without a discipline. For instance, they are concerned with the impact upon science of the governmental, economic, military, and religious institutions in society. To them, a discipline is rather open to the ideas from other disciplines. Instead of seeing a discipline as a collection of well defined research areas (the internalist view), it is seen as a complex network of interacting researchers and its research ideas are seldom viewed as originating from one specific research area.

Both models are based on the assumption that juxtaposition of ideas from different fields may change the existing paradigm or produce a new paradigm, and without such a scientific paradigm no discipline is able to produce systematic and therefore cumulative body of knowledge. If so, information retrieval and librarianship cannot be exceptions. Which disciplines influence the forumulation of the paradigmatic content of information retrieval and librarianship? To what extent is information retrieval and librarianship open or closed to the ideas of other disciplines or, perhaps more importantly, to each other?

It appears that citation analysis has not been used to investigate the interface between admittedly closely related activities such as librarianship and information retrieval. Given this traditional affinity we would expect that researchers in both would depend heavily on each other. Furthermore, they would both depend on the literature of an homogeneous group of other disciplines. Since information retrieval is the younger of the two we would also expect that it would be more open to the ideas of other disciplines than librarianship.

IV. Method

Sample and Source of Data

A total of 1159 citations were obtained from the 12 most cited journals in information retrieval and librarianship. This set of journals was obtained in the following way. First, Ulrich's International Periodical Directory, 18th Edition (New York, Bowker, 1979) was consulted and a list of journals under the heading of Library and Information Sciences was compiled. This list was then checked against the Social Science Citation Index (SSCI), Journal Citation Reports, Ranking Package (Philadelphia, Institute for Scientific Information, 1979) to determine the library and information retrieval journals with the highest "impact factors." The impact factor is a measure of the number of articles in a particular journal cited by other journals divided by the total citations from all journals in SSCI.

From this list, 6 journals with the highest impact factor in information retrieval were identified. Similarly, 6 were identified in librarianship. It was decided to take sample citations from the last complete year

(1978, in most cases). Up to five articles from each journal were randomly chosen. All of the references from each of those articles were coded. The source journals and number of references are shown in Table 1.

Table 1

SOURCE JOURNALS AND SAMPLE SIZE

Information Retrieval	No. of References	Librarianship	No. of References
Jl of Documentation	232	Library Journal	51
Jl of American Society for Info. Sci.	98	Lib. Res. & Tech. Serv.	244
Jl of Chemical Info. & Comp. Sci.	25	Special Libraries	33
Information Sciences	44	Library Quarterly	107
Information Processing & Management	46	RQ	60
Information Reports & Bibliography	21	Library Trends	198

N = 1159

Measures

Each citation in the sample was classified as to which source specialty it came from (i.e., librarianship or information retrieval). Each cited item was classified by its parent discipline, using Ulrich's for journals and the researchers' judgements for monographs. In all, 26 different disciplines were identified as contributing to our sample literature.

From this data we developed the following measures:

1. From-Within Measure (Self-citation). In order to measure the impact exerted from within a specialty, the percentage of self-citation was calculated , i.e., the frequency with which a specialty cites itself. For each source specialty (librarianship and information retrieval), we calculated the proportion of citations to journals in that specialty.

2. From-Without Measures. To determine the relative influences exerted from outside, the frequency of citations coming from different disciplines was tabulated. This resulted in (a) a cross-fertilization measure, and (b) an openness measure.

   (a) Cross-fertilization. By cross-fertilization we mean the proportion of citations made by one to the other source specialty. In other words, the extent to which librarianship cited information retrieval journals and vice versa. It implies the degree to which the two specialities draw from each other.
   (b) Openness. This is a measure which indicates the degree to

which our source specialties are open to the ideas of other disciplines. It equals the total number of citations from a particular discipline divided by the total citations from all other disciplines.

V. Findings and Discussion

Table 2 shows the citation data by the 15 most cited disciplines. Numbers in the table are the proportions of citations from various disciplines, including self-citations.

Of the 1159 citations, 506 were from monographs. Information retrieval journals cited monographs 47% of the time while librarianship journals cited monographs 42% of the time. In light of the recency of research interest of information retrieval, 47% citation rate to monographs is interesting.

From-Within (Self-Citation) Measure

As is evident from Table 2 about half of all citations made by the source specialties were self-citations (information retieval = 48%; librarianship = 48%). Although the self-citation rates appear to be high, they are not as high as the 70% self-citation attributed to science as a whole nor to the rates found for mathematics (79%), Physics (79%), Engineering (70%), and other science disciplines (Earl, et.al., 1969). Although such data are not available, it would be interesting to compare the self-citation rates of our source specialties with those of other social science disciplines.

From-Without Measures

It is also evident that the two source specialties did not cite each other to any great extent. In other words, cross-fertilization between the two, in spite of their historical affinity, was minimal. Evidence of a dialog between these two specialties is not found in the literature analysis of this study. Librarianship cited information retrieval journals only 9% of the time and information retrieval cited librarianship only 7% of the time.

The finding of self-citation and cross fertilization seem to suggest the following. That is, given the 70% self-citation rate of science as a whole, our source specialties have significantly lower self-citation rates. This would seem to reject the internalist argument. Self-citation rate is not significantly high enough to conclude that the impact on the scientific paradigm originates largely from within.

Table 2 also shows the other disciplines which were frequently cited by the two source specialties. Overall, information retrieval drew from other disciplines (excluding librarianship) about 45% of the time and librarianship drew from other disciplines (excluding information retrieval) about 43% of the time. It appears that both specialties are equally open to the ideas of other disciplines. However, the nature of impact of other disciplines upon the source specialties is different. It can be seen that information retrieval borrowed most heavily from linguistics, computer science, mathematics, and general science/current events journals; librarianship borrowed from psychology, political science and photography/

micrography.

We conclude, therefore, that although the total citations from other disciplines were almost 50%, our source specialties drew from a different set of disciplines.

Table 2

PROPORTIONS OF DISCIPLINARY CONTRIBUTIONS TO THE LITERATURE OF INFORMATION RETRIEVAL AND LIBRARIANSHIP

Cited Discipline	Citing (Source) Specialty: Information Retrieval	Librarianship
Information Retrieval	.48 (1)	.09 (2)
Librarianship	.07 (2)	.48 (1)
Education	.03	.03
Management	---	.02
Computer Science	.09	---
Linguistics	.10	---
Sociology	.02	.03
Political Science/Law/Government	.02	.06
Mathematics/Statistics	.05	.01
Engineering	.02	---
Photography/Micrography	---	.05
Psychology	.07	.09
General (3)	.07	.03
Others (4)	.01	.05
Unknown	.03	.06

(1) Rate of self-citation
(2) Rate of cross-fertilization
(3) Includes journals on general science and/or current events such as Science, Nature, Time, etc.
(4) All other disciplines with less than .01 citation rate each.

VI. Implications and Suggestions for Further Studies

The fact that our sample literature drew from at least 26 different disciplines may imply that the field of librarianship and information retrieval are essentially interdisciplinary in nature. Their paradigmatic content may be determined by various problems which they face and feel compelled to solve. Required knowledge for problem-solution may be drawn from any pertinent disciplines as the need arises rather than as a result of internal intellectual, systematic development. If this is true then librarianship-information retrieval may be in the pre-paradigmatic stage of development.

Three extensions of the present study appear interesting: (1) investigation of citation patterns over time to determine if disciplinary contributions change; (2) investigation of frequently cited authors to determine their disciplinary background and respective contribution to the field; and, (3) interviewing the notable authors and conducting a content analysis of the literature of our field to determine its topical emphasis, research methodologies, and research priorities with a view toward characterizing

how or whether the field is developing into a science with an identifiable paradigm.

REFERENCES

American Institute of Physics. Information Division. A Program for a National Information System for Physics and Astronomy 1971-75. 1970.

Borenius, G. and S. Schwarz. "Remarks on the Use of Citation Data in Predictive Models of Scientific Activity." Information Storage and Retrieval. 8:4 (August 1972), pp. 171-175.

Burton, R. E. and R. W. Kebler. "The 'Half-life' of Some Scientific and Technical Literatures." American Documentation. 11:18-22.

Cole, Jonathan and Stephen Cole. Social Stratification in Science. Chicago, University of Chicago Press, 1973.

Crane, Diana. Invisible Colleges: Diffusion of Knowledge in Scientific Communities. Chicago, University of Chicago Press, 1972.

Earle, P. and B. C. Vickery. "Subject Relations in Science/Technology Literature." Aslib Proceedings. 21 (1969), pp. 237-43.

Garfield, Eugene. "Citation Analysis as a Tool in Journal Evaluation." Science. 178 (November 1972), pp. 471-479.

Gross, P.L.K. and E.M. Gross. "College Libraries and Chemical Education." Science. 66 (October 1927), pp. 385-89.

Hafner, Arthur W. "Primary Journal Selection Using Citations from an Indexing Service Journal; a Method and Example from Nursing Literature." Bulletin of the Medical Library Association. 64(4) (October 1976), pp. 392-401.

Holton, G. "Scientific Research and Scholarship: Notes toward the Design of Proper Scales." Daedalus. 91 (1962), pp. 362-99.

Martyn, John. "Citation Analysis." Journal of Documentation. 31(4) (December 1975), pp. 290-97.

Meadows, A.J. "The Citation Characteristics of Astronomical Research Literature." Journal of Documentation. 23 (1967), pp. 28-33.

Meadows, A.J. Communication in Science. London, Butterworths, 1974.

Mekhtiev, A.M.; Arakelov, R.K.; Mellion, S.P.; and Kiel Gavina. "Using an Index of Cited Literature to Analyze the Current State and Trends of Information Science." Automatic Documentation and Mathematical Linguistics. 11(1) (1977), pp. 49-59.

Parker, E.B., Paisley, W.J. and R. Garrett. Bibliographic Citations as Unobtrusive Measures of Scientific Communication. Stanford, Calif., Stanford University Institute for Communication Research. 1967.

Price, D.J. "Networks of Scientific Papers." Science. 149 (1965), pp. 510-15.

Scales, Pauline A. "Citation Analysis as Indicators of the Use of Serials; A Comparison of Ranked Title Lists Produced by Citation Counting and from Use Data." Journal of Documentation. 32(1) (March 1976), pp. 17-25.

Small, Henry. "Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents." Journal of the American Society for Information Science. 24(4) (July/August 1973), pp. 265-69.

Subramanyan, K. "Citation for Journal Selection." Special Libraries. 66 (August 1975), pp. 367-71.

Swanson, D.R. "Scientific Journals and Information Services of the Future." American Psychologist. 21 (1966), pp. 1005-10.

Virgo, Julie A. "A Statistical Procedure for Evaluating the Importance of Scientific Papers." Library Quarterly. 47(4) (October 1977), pp. 415-30.

Weiss, P. "Knowledge: A Growth Process." Science. 131 (1960), pp. 1716-19.

Windsor, Donald A. "Rational Selection of Primary Journals for a Biomedical Research Library; The Use of Secondary Journal Citations." Special Libraries. 64 (October 1973), pp. 446-51.